

University of Groningen

Introduction

Kuipers, Theo A.F.

Published in:
EPRINTS-BOOK-TITLE

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2007

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Kuipers, T. A. F. (2007). Introduction: Explication in philosophy of science. In *EPRINTS-BOOK-TITLE* (pp. vii-xxii). s.n..

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

INTRODUCTION. EXPLICATION IN PHILOSOPHY OF SCIENCE

Theo A. F. Kuipers

1 AN IMPORTANT, THOUGH LARGELY IMPLICIT, METHOD IN PHILOSOPHY OF SCIENCE

Scientists use concepts, principles and intuitions that are partly specific for their subject matter, but they also share part of them with colleagues working in different fields. Compare, for example, the chemical notion of the ‘valence’ of an atom with the notion of ‘confirmation’ of a hypothesis by certain evidence. An important task of ‘the philosophy of (the special science of) chemistry’ is the explication of the concept of ‘valence’. Similarly, an important task of ‘general philosophy of science’ is the explication of the concept of ‘confirmation’. In both cases it is evident that this only makes sense if one tries to do justice, as much as possible, to the actual use of these notions by scientists, without however following this use slavishly. That is, occasionally a philosopher may have good reasons for suggesting to scientists that they should deviate from a standard use. Frequently, this amounts to a plea for differentiation in order to stop debates at cross-purposes due to the conflation of different meanings.

What has been said about concepts, also applies to principles and intuitions of scientists, which may or may not be paradoxical. Compare the subject specific ‘principle of the conservation of energy’ and the general intuition of ‘diminishing returns from repeated tests’. Both aren’t crystal clear, they need explication; of course, beginning with the explication of the concepts involved.

Although the term ‘explication’ is not often used by philosophers, it is clear that when they discuss the meaning of concepts and propose or report specific definitions, characterizations, models, theories, accounts, conceptions, (rational) reconstructions or formalizations of them, they are practicing concept explication in a more or less explicit and rigorous way. Similarly, when philosophers propose or report specific analyses, accounts, reconstructions or formalizations of principles and intuitions, or dissolutions of paradoxes, explication is at stake. Both kinds of activity belong to the dominant ones among (systematic, constructive, analytical) philosophers, and not in the least among philosophers of science. However, explicitly calling these activities ‘explication’ is not very popular, let alone using the explication terminology in presenting results. There seem to be at least three reasons for the reluctance to use the word ‘explication’. First, the word itself may

be found a bit too affected. Second, making the application of the method explicit may not only lead to rather cumbersome texts, but also appear to be a difficult task. Finally, many philosophers do not like to be associated with the logical empiricists that introduced ‘(concept) explication’ around 1950 as a technical term for this philosophical method, viz. Rudolf Carnap and Carl Hempel.

This general volume of the *Handbook of the Philosophy of Science* deals with focal issues of a general nature, whereas the special volumes address topics relative to a specific discipline. Each volume contains several contributions that illustrate (largely implicitly, as suggested) the use within philosophy of science of the *method of explication* in one form or another. In a sense, it is what is left of the undoubtedly more rigorous ‘logical analysis’ with which philosophy of science started in the first half of the 20th century and which had to give way to the ‘historical approach’ (see Aliseda and Gillies, this volume).

Below I will first describe and further articulate the method of explication, paying particular attention to evaluation reports of proposed explications, to the principled possibility of progress, and to explication starting with idealization, followed by successive concretization. Then I will give a survey of the ten chapters in this volume, with emphasis on the most important concepts and intuitions of which the explication is discussed.

2 EXPLICATION OF CONCEPTS AND INTUITIONS IN SCIENCE

The method of concept explication

As explained above, most philosophers apply the method of explication only informally and implicitly. Hence, the reconstruction of this method is itself a kind of meta-explication. Notably Rudolf Carnap [1950, 1966] and Carl Hempel [1950/1966; 1952], but also John Kemeny and Paul Oppenheim [1952] have articulated the method. Here I will freely describe and develop the method in their spirit. As suggested, it has been and can be applied in various degrees of rigor and explicitness.

The point of departure of concept explication in general philosophy of science is an informal, intuitive concept, the *explicandum*, which is frequently used by scientists in different fields. The aim is to define a concept, the *explicatum*, satisfying three desiderata for concept formation in general and some specific ones. The three *general* desiderata are *precision*, *fruitfulness* and *simplicity*. The *specific* desiderata pertain to the *similarity* with the informal concept. This similarity is to be evaluated by two kinds of criteria:

1. The explicatum should apply to evident, undisputed (*types of*) *examples* of the informal concept and it should not apply to evident, undisputed (*types of*) *non-examples*.
2. The explicatum should fulfil (other) *conditions of adequacy* that have been derived from the informal concept, and occasionally it should violate *con-*

ditions of inadequacy. Evident (non-)examples may be generalizable to a condition of (in)adequacy.

Due to the general and specific desiderata, concept explication is an active enterprise, as opposed to the passive cognate of concept or meaning *analysis* (and its precursor ‘logical analysis’). In the latter activity one merely tries to do justice to the uses of an informal concept, hence without any further desiderata. To be sure, meaning analysis may be very useful, if only to search for evident examples and non-examples and for finding further conditions of (in)adequacy for concept explication.

Confirmation

Let me briefly illustrate the specific desiderata with the famous example of the explication of ‘confirmation’. It is generally assumed, at least *prima facie*, that “this entity is a black raven” confirms the hypothesis that “all ravens are black”, and hence provides an evident example of a confirmation. This is so-called Nicod’s criterion. On the other hand, a white raven is a counterexample and hence an evident non-example. A black or non-black non-raven, e.g. a black or white tie, also seem evident non-examples. However, it also seems plausible that the explicatum of confirmation should satisfy the ‘equivalence condition’, that is, if certain evidence confirms a certain hypothesis then it confirms any of its logically equivalent versions. From this it would follow that a non-black non-raven, e.g. a white tie, also confirms “all ravens are black”, for it is, according to Nicod’s criterion, an evident confirming instance of its logically equivalent version “all non-black objects are non-ravens”. This is Hempel’s so-called paradox of confirmation ([Hempel, 1945/1965]; see also both Ladyman and Niiniluoto, this volume). It is clear that one can accept this consequence, assuming perhaps certain test conditions, or one can start to dispute either what is to be counted as an evident example or how plausible the equivalence condition in fact is.

Following Hempel [1945/1965], scholars accepting the equivalence condition have also argued that the explicatum of confirmation should satisfy the ‘(special) consequence condition’, that is, if certain evidence confirms a certain hypothesis then it confirms any of its consequences. However, it may also be argued that the opposite condition should be taken as a condition of adequacy, that is, if certain evidence confirms a certain hypothesis then it confirms any stronger hypothesis, the so-called ‘converse consequence condition’. Although both conditions may have some intuitive appeal, as Hempel already pointed out, it is easily seen that imposing them together leads to absurd conclusions: any evidence confirms any hypothesis as soon as it confirms some hypothesis (if E confirms H , it confirms $H \& H^*$, by the converse consequence condition, hence it confirms H^* by the consequence condition). For Hempel this was a reason to (implicitly) classify the converse consequence condition as a condition of *inadequacy*.

General desiderata

As the example of ‘confirmation’ illustrates, it may not at all be an easy matter to agree upon the desiderata for similarity. This is not merely a problem among philosophers of science. The claim is that scientists use many concepts of which it is far from clear that they are used in a consistent way. At the same time, both scientists and philosophers of science may have the feeling that one or more specifications of the concept will make it clear that the use of the concept within science is in fact not as confused as it might seem. In this way it has not only become clear that concept explication is far from a straightforward task, but the general desiderata for explication proposals have also come into the picture. I will briefly discuss them.

1. *Precision (or exactness)*. This desideratum is far from precise itself, not in the least because it depends very much on the type of concept one wants to explicate and the purposes that one has primarily in mind with this task. In other words, precision is required as far as relevant and possible. However, some aspects can be put forward as desirable in general. The explicatum should neither lead to inconsistent nor to trivial conclusions. It is also plausible to require syntactical and semantical determinateness. In particular, it should be clear to which syntactic category the relevant term belongs: is it a property term, a relational term or a function term? Moreover, is it classificatory (E confirms H, or not), comparative (E confirms H more than H*) or quantitative (the degree of confirmation of H by E is c)? Note that these divisions are incomplete and do not coincide: for example, an equivalence relation is not comparative and not every function is quantitative. Finally, to reach a sufficient degree of precision it is often desirable to introduce at least some kind of formalization.
2. *Fruitfulness*. Besides minimally avoiding trivialization, already captured by the precision desideratum, the proposal should be fruitful in one or more of the following senses. It should throw new light on the problems that motivated the explication enterprise in the first place. This may concern the explication of a puzzling intuition or the dissolving of a paradox. As far as reaching this goal is considered as a condition of adequacy this will be taken into account in the specific evaluation. However, fruitfulness may also concern the desire for a coherent account of related concepts, like providing the solution of a jigsaw puzzle. As a matter of fact, the point of departure of explication frequently is a network of informal concepts, one of which is chosen as the primary target. For example, the concepts of ‘confirmation’ and ‘induction’ are related, and it may be fruitful to aim at an explication of ‘confirmation’ that can be seen as a weak kind of ‘induction’. Besides such intended successes, an explication is even more fruitful when it also leads to unintended, extra successes.
3. *Simplicity*. The concept of simplicity is itself an example of a concept one

has tried to explicate through the ages. In particular, the question what scientists mean when they prefer one explanation or one theory to another because the former is claimed to be simpler than the latter. Although various attempts have been made, without leading to a generally accepted answer, there has certainly been made progress, in particular as far as quantitative theories are concerned (see, [Sober, 1998]).

Of course, philosophers (and scientists) may disagree about how important these general desiderata are, and I would not like to take a stance in this. However, one thing is clear: when a proposal satisfies one of these desiderata in a striking way, in particular in comparison with a rivaling explication, it is almost generally conceived as a positive point, other things being equal.

Evaluation report

The evaluation report of an explication makes clear to what extent the proposed explicatum satisfies the general and specific desiderata. As to the three general desiderata, an informal scoring will usually be the best that can be obtained. Of course, for the question whether progress has been made, see below, comparative scoring is more important than separate scoring.

As to the specific desiderata, it might be wise to use in the evaluation report the problem solving terminology of successes and problems that has been introduced by Thomas Kuhn [1962] and Larry Laudan [1977] for the evaluation of scientific theories.

Starting with evident examples and non-examples, on the *success side* of the report we have the ‘true positives’ and the ‘true negatives’, that is, the evident examples fitting into the proposed explication and the evident non-examples excluded by it, respectively. On the *problem side* we get the ‘false positives’ and the ‘false negatives’, that is, the evident non-examples that nevertheless fit and the evident examples that don’t fit, respectively. False positives show that the explication is too wide and false negatives that it is too narrow.

Besides clear-cut classifications of some evident (non-)examples as problems, there may well have been identified other ‘evident’ (non-)examples of which the question has arisen whether they are really evident. Such disputed (non-)examples may even lead to revision of the connotation of the concept. That is, formerly evident examples and non-examples may become non-examples and examples, respectively, though not necessarily evident ones. They may be classified as (non-evident or) questioned examples and non-examples, respectively. Of course, other (non-)examples about which there were and are no strong feelings may also be classified as questioned. The suggested type of connotation revision need not occur when a problem, a false positive or false negative, can easily be seen as resulting from the fact that the proposal is still very idealized or naïve and, hence, that refinement or concretization of the proposal is on the agenda. I will come back to this.

Regarding the other type of specific desiderata, viz., conditions of (in)adequacy, a similar terminology is called for. Fulfilled conditions of inadequacy and violated

conditions of inadequacy are successes, and violated conditions of adequacy and fulfilled conditions of inadequacy are problems. Violated conditions of adequacy show that the explication is too wide and fulfilled conditions of inadequacy that it is too narrow. Again, conditions may be classified as questioned. Unfulfilled conditions of adequacy and fulfilled conditions of inadequacy may become disputed and lead to connotation revision, but such problems may again also be seen as due to initial idealization, requiring concretization of the explication in order to become (un)fulfilled.

In sum, restricted to specific desiderata, an evaluation report of a proposed explication minimally specifies on the success side evident examples that are covered and conditions of adequacy that are fulfilled. On the problem side it at least lists evident examples that are not covered and conditions of adequacy that are not fulfilled. Finally, on both sides there may occur old and new (non-)examples and conditions of which the status still is or has become questioned.

Let me specify part of Hempel’s implicit evaluation report for the ‘prediction criterion’ [Hempel, 1945/1965] or ‘conditional deductive confirmation’ [Kuipers, 2000] as an explication proposal for confirmation: E confirms H assuming condition C if and only if $H \& C$ logically entails E , but C alone does not, where C and H are non-contradictory and E even contingent. I insert question marks to indicate (non-)examples and conditions of (in)adequacy that have been disputed by others as evident or desirable, respectively. I add the ‘exclusion of something like probabilistic confirmation’ as a plausible condition of *inadequacy*, and hence as a problem, to be solved by subsequent concretization.

<i>Specific desiderata</i>	<i>Successes</i>	<i>Problems</i>
<i>Evident (non-)examples</i>		
Evident examples, relative to “All ravens are black”	True positive: - a black raven	False negative:
Evident <i>non</i> -examples, relative to “All ravens are black”	True negative: - a white raven - a black tie (?)	False positive: - a white non-raven (?)
<i>Conditions of (in)adequacy</i>		
Conditions of adequacy	Fulfilled: - Nicod’s criterion - equivalence condition	Unfulfilled: - consequence condition (?)
Conditions of <i>inadequacy</i>	Unfulfilled:	Fulfilled: - converse consequence condition (?) - excluding probabilistic confirmation

Although it is intuitively appealing to continue to take evident non-examples and conditions of inadequacy explicitly into account in the evaluation report, it will be clear that they can implicitly be taken care of by appropriate conditions of adequacy, excluding the evident non-examples, and by appropriate evident examples, breaking the conditions of inadequacy, respectively. For instance, counterexamples to a general hypothesis are excluded as evident non-examples of confirmation by the plausible requirement that confirming evidence at least has to be compatible with the hypothesis. A well known example of the second type is the drawback of Popper's explication of 'closer to the truth', which happened to leave only room for *true* theories closer to the truth than other theories (see Niiniluoto, this volume). This feature is an evident condition of inadequacy, that can be broken by requiring to leave room for *the possibility* that the theory of Einstein is closer to the truth than the theory of Newton, even if the former would turn out to be false.

It is important to note that conditions of adequacy for concept explication in general may be of a formal or an empirical nature, resulting from previous meaning analysis or from previous empirical analysis, respectively [Hempel, 1952]. An example that combines both kinds of conditions is Einstein's explication of simultaneity of events at a distance. It was guided by the need to obey the empirical law of the constancy of the velocity of light in all frames of reference and by the convention to average that speed over outgoing and return paths. As this law is a cornerstone of Einstein's theory of relativity, the example illustrates that concept explication may be strongly laden with empirical laws, theories and conventions. Of course, conditions of adequacy (and evident examples) based on empirical laws and theories may become questioned when they come under attack or when merely different, but compatible perspectives are possible. For example, the empirical notion of 'absolute temperature' is strongly based on some (asymptotic) empirical laws, whereas the theoretical notion of absolute temperature is heavily based on the theory of thermodynamics, and for this reason also is called 'thermodynamic temperature'.

Formal conditions of adequacy may be of a justificatory nature, for example in the sense that the explication should provide some required existence and uniqueness conditions. An important historical example is that the explications (more specifically, the resulting definitions) of the notion of limits of mathematical sequences and series have to satisfy such conditions, at least as much as possible. Formal conditions may also pertain to the intended justification of a certain intuition, for example, the explication of the 'severity' of a test should realize the intuition of 'diminishing returns from repeated tests', see below. Finally, they may pertain to dissolving a paradox, for example Hempel's paradox of confirmation as resulting from a first explication attempt, see above.

Progress in concept explication

This brings us to the comparative evaluation of two explications of the same concept and the possibility of progress in concept explication or, briefly, *conceptual*

progress. Concept explication frequently leads to the conclusion that there are in fact two or more concepts that have to be distinguished, leading to a branching of the explication. However, when two explications of a concept are considered as rivals, the plausible question arises whether the one is better than the other. It is easy to see that there is a possibility to define ‘strictly better’, which however assumes agreement about the relevant evaluation reports, both with respect to the relevant items as to their scores. The latter is the easiest to imagine by one person and, if relevant, from one theoretical perspective. Moreover, it explains why another person frequently disputes the claim by one scholar of having made progress on the basis of a disagreement about the separate evaluation reports. But, as suggested, this does not prevent the following definition, neglecting evident non-examples and conditions of inadequacy, for the reasons indicated above.

Definition. *E2* is a *strictly better* explication of a concept than *E1* if and only if:

1. *E2* satisfies the general desiderata at least as well as *E1*
2. *E1* and *E2* share all questioned examples and conditions of adequacy
3. *E2* covers all evident examples covered by *E1*
4. *E2* fulfils all conditions of adequacy fulfilled by *E1*
5. *E2* covers some more evident examples and/or fulfils some more conditions of adequacy

It is easy to imagine stronger, weaker and refined versions of various kinds, for example, by requiring unintended successes, by merely counting numbers of covered examples and fulfilled conditions, and by assigning weights to them, respectively. Moreover, changes of questioned examples and conditions may be taken into account. In the terms of Rawls, further articulated by Thagard [1988], one may summarize the overall conclusion that one explication is better than another by claiming that the first better approaches a *reflective equilibrium* than the second. However, the most important point is that the ideal of conceptual progress can be defined in a strict sense and hence that it can function as a regulative idea.

The notion of an evaluation report and the definition of conceptual progress are modeled along the lines for evaluating theories and defining empirical progress in terms of counterexamples and explained empirical laws [Kuipers, 2000, Chs 5/6; Kuipers, 2001, Chs 7/8]. Hence, the partial, formal analogy between empirical and conceptual progress need not be surprising. Another reason why such an analogy may be expected is that in both cases we deal with progress in problem solving, albeit of different kinds.

An idealized start followed by successive concretization

As suggested before, a first explication may be a highly idealized way of catching cases and conditions, with the explicit intention to set up subsequent more realis-

tic explications by accounting for cases and aspects that have first been neglected. This is a useful strategy, not only for concept explication, but also for concept formation in general. Moreover, it is largely analogous to what has been explicated for the empirical sciences as the strategy of idealization and concretization or factualization [Nowak, 1980; Krajewski, 1977]; see also [Kuipers, 2002/to appear]. Besides being a strictly better explication, the specific criterion for being a successful concretization of an idealized one is that the latter is an extreme special case of the former.

Well known examples in the sphere of confirmation are, leaving out some technical restrictions, the transition from ‘deductive confirmation’ (hypothesis H logically entails evidence E) to, on the one hand, ‘conditional deductive confirmation’ (assuming condition C , H entails E), used above for illustrating the notion of an evaluation report, and to ‘probabilistic confirmation’ ($p(H/E) > p(H)$), on the other. In the first example of concretization, we get the original back in the extreme special case that the condition C is a tautology, and in the second case when $p(E/H) = 1$. Of course, both types of concretization can be combined (see Niiniluoto, this volume, for further details on confirmation).

Another example of concretization is the transition from simple or Bayesian conditionalization to ‘Jeffrey conditionalization’, taking into account that the posterior probability of a hypothesis may be based on evidence about which one is not certain. I just quote from the *Stanford Encyclopedia of Philosophy* [Joyce, 2003, 13–4]:

Simple Conditioning:

If a person with a ‘prior’ such that $0 < \mathbf{P}(E) < 1$ has a learning experience whose sole immediate effect is to raise her subjective probability for E to 1, then her post-learning ‘posterior’ for any proposition H should be $\mathbf{Q}(H) = \mathbf{P}_E(H)$. [Here $\mathbf{P}_E(H)$ is standardly defined as $\mathbf{P}(H \& E)/\mathbf{P}(E)$, TK]

[...]

Jeffrey Conditioning

If a person with a prior such that $0 < \mathbf{P}(E) < 1$ has a learning experience whose sole immediate effect is to change her subjective probability for E to q , then her post-learning posterior for any H should be $\mathbf{Q}(H) = q\mathbf{P}_E(H) + (1 - q)\mathbf{P}_{\neg E}(H)$.

Obviously, Jeffrey conditioning reduces to simple conditioning when $q = 1$. [That is, the latter is an extreme special case of the former, TK]

Explication of intuitions and dissolving paradoxes

Turning briefly to the explication of intuitions and principles and the dissolving of paradoxes, the first task is of course the explication of concepts that are crucial

for the formulation of the intuition or paradox. In case of intuition explication the subsequent task is to prove a theorem to the effect that the intuition, if reformulated in explicated terms, becomes justified, demystified or undermined, whatever the case may be. In case of dissolving a paradox, it has to be shown that it can no longer be construed in the explicated terms.

One example is the quantitative and qualitative explication of the intuition that empirical progress is functional for truth approximation, [Niiniluoto, 1987; Kuipers, 2000], respectively; see also Niiniluoto, this volume).

Another example of intuition explication is the ‘diminishing returns’- intuition. For whatever my explication of it is worth, it has all indicated features and is easy to present in a short paragraph. Popper expressed the intuition as follows: “There is something like a law of diminishing returns from repeated tests” [Popper, 1963, p. 240]. It is the idea that the returns, in the sense of the relevance, impact or severity of repeated tests decreases in one way or another. The core idea of my ‘non-inductive’ explication [Kuipers, 1982] is the following. Explicate ‘the returns of n repeated tests’ as their a priori objective severity: that is, the prior probability that they will generate at least one counterexample of the relevant generalization. Assuming random tests, with fixed probability $(1 - q)$ of a counterexample, this severity is $1 - q^n$. Hence, the extra severity of an extra test, the $n + 1$ -th test, is $(1 - q^{n+1}) - (1 - q^n) = q^n(1 - q)$. Note that this equals the prior probability that the $n + 1$ -th test generates the *first* counterexample. Now it is easy to prove that this probability diminishes to 0, assuming that $q < 1$. The last assumption implies that the generalization under test is false, which is according to this explication an underlying assumption of the intuition.

Concluding remarks

Ever since logical empiricists like Carnap, Hempel and Kemeny and Oppenheim presented their idea of concept explication in theory and practice, it has become an important, if not dominant, method in analytical philosophy of science. However, presumably because it was considered to be self-evident, philosophers have internalized the method and its applications usually remain largely implicit. As said, there seems to be moreover a certain reluctance to make the method explicit, for several reasons. Recall that one important reason was that it is not always easy to discover, disentangle and classify the specific desiderata. To identify these desiderata in a text, one may start with searching for objections that are claimed to have been met or have to be met according to the author. For this purpose the above-presented apparatus may be helpful for close reading philosophical texts and for doing philosophical research, although it should of course not be used as a Procrustean bed. In the present context of a series of handbooks in philosophy of science, hopefully also with many non-philosophers as readers, for whom the method will be less self-evident, it may also be useful to have spelled it out. Moreover, below it will be used as a selective point of view in presenting a survey of the chapters to come in this general volume.

3 SURVEY OF CHAPTERS

Chapter 1: Laws, theories, and research programs (Theo Kuipers)

Observational laws, theories and research programs are three of the main units in empirical science. The informal distinction between observational (or experimental) laws and proper theories is illustrated with relevant examples and shown to be of crucial importance for the short and long term dynamics of science. However, it is difficult to explicate the two concepts as strictly distinct, mainly because there is no theory-free observation, which suggests a merely gradual distinction. Fortunately, as notably Hempel and Sneed have put forward, a theory-relative explication of the two concepts can save the distinction.

As Kuhn and Lakatos have shown, the concept of paradigm or research program can capture the short and long term development of science as sequences of theories that are bound together by core ideas that are not given up. This strategy is possible due to unavoidable, but revisable auxiliary hypotheses. However, explicating the notion of a research program not only requires branching in at least four kinds of programs, viz., descriptive, explanatory, design and explicative ones, it also requires the introduction of some degrees of strength. Together they yield the view on intra- or interdisciplinary interaction in science as a matter of competing or cooperating research programs.

Since the heydays of the logical empiricists, philosophers of science have felt that the notion of an empirical theory can be explicated such that its structure becomes clear. The structuralist explication of theories, due to Suppes and Sneed, is presented by way of an optional intermezzo. It can be used for reconstructing theories, which have at least some formal aspects, in order to further study their reach and developmental potentials.

Chapter 2: Past and contemporary perspectives on explanation (Stathis Psillos)

The term ‘explanation’ is perhaps one of the most commonly used terms by scientists and philosophers of science. The explicative endeavors in the past and the present have not only shown that some distinctions have to be introduced, e.g., between deductive, statistical, teleological, and historical explanation, but also that some basic conceptions have fundamentally remained the same. For some scholars the changes are a matter of conceptual progress, for others they are merely a matter of changing emphasis and interests. The first part of the chapter describes how some major thinkers, from Aristotle, through to Descartes, Leibniz, Newton, Hume and Kant, to Mill, conceived of explanation. The second part offers a systematic examination of the most significant and controversial contemporary models of explanation.

The first part opens with Aristotle’s conception — the thought that explanation consists in finding out *why* something happened and that answering why-questions

requires finding causes — which has set the agenda for almost all subsequent thinking about explanation. This part further discusses the search for a coherent account of laws of nature, as opposed to mere generalizations, causation and explanation in the thought of the early modern philosophers. This culminates in John Stuart Mill's first well-worked out model of scientific explanation, which was based on the idea that there is no necessity in nature and that, ultimately, explanation amounts to unification into a comprehensive deductive system, whose axioms capture the fundamental laws of nature.

The second part starts with the logical empiricists' attempt to explicate and legitimize the concept of causation by subsuming it under the concept of a deductive-nomological explanation, that is, under the concept of explanation as explicated by Hempel and his followers. It moves on to discuss the re-appearance of genuinely causal models of explanation as well as the re-appearance and development of the Millian idea that explanation amounts to *unification*. Next to this it pays attention to models of statistical and probabilistic explanation, with or without the claim to capture causal forms of such explanations. It ends with an examination of historical and teleological approaches to explanation.

Chapter 3: Evaluation of theories (Ilkka Niiniluoto)

In assessing the cognitive merits of science, one might take a hypothetical theory as the basic unit of evaluation. The traditional virtues of a good theory include consistency, truth, prior and posterior probability, information content, empirical content, explanatory and predictive power, problem-solving capacity, simplicity, accuracy, approximate truth, and truthlikeness. All these notions have been subjects of explication. In this chapter some of the most important explications of most of them are discussed.

The chapter further discusses qualitative, comparative, and quantitative explications of the confirmation of theories by means of available evidence. The results formulate conditions under which empirical success inductively supports the truth or truthlikeness of a theory. Particular attention is given to abductive confirmation due to successful deductive or inductive explanation and prediction. Conditions for accepting a theory as true or truthlike are also discussed. For example, attention is paid to the explication of so-called 'inference to the best explanation' and the explication of the intuition of convergence to the truth.

Chapter 4: The role of experiments in the natural sciences. Examples from physics and biology (Allan Franklin)

Whereas the concept of experiment is not very controversial, at least not in the natural sciences, the epistemological role of experiments is controversial, partly because experiments play many roles. One of their important roles is to test theories and to provide the basis for scientific knowledge. They can also call for a new theory, either by showing that an accepted theory is incorrect, or by exhibiting

a new phenomenon which needs explanation. Experiment can provide hints toward the structure or mathematical form of a theory and it can provide evidence for the existence of the entities involved in our theories. It can also measure quantities that theory tells us are important. Finally, it may also have a life of its own, independent of theory. Scientists may experimentally investigate a phenomenon just because it looks interesting. This will also provide evidence for a future theory to explain.

In all of this activity, however, we must remember that science is fallible. Theoretical calculations, experimental results, or the comparison between experiment and theory may all be wrong. If experiment is to play the indicated important roles in science then we must have good reasons to believe experimental results. The chapter presents an epistemology of experiment, that is, a set of strategies that provides reasonable belief in experimental results. Scientific knowledge can then be reasonably based on these experimental results. The view is defended that nature, as revealed by experiment, plays an important and legitimate role in science. The examples come, primarily, although not exclusively, from physics, but these episodes seem typical of the natural sciences. Several examples from biology are also included. These examples do not only provide evident cases for the branched explication of the nature and role of experiments, but also evident cases for the explication of notions like confirmation and refutation.

Chapter 5: The role of experiments in the social sciences. The case of economics (Wenceslao Gonzalez)

In the social sciences the notion of an experiment is less clear than in the natural sciences. This chapter discusses the traditional conception of experiments in the social sciences, and its more recent, enlarged vision. Like in the natural sciences, the analysis of the role of experiments is among the central topics in the methodology of the social sciences. However, doing experiments in the social sciences, in general, and in economics, in particular, has not always been accepted, and it is still an issue that raises objections. The case of economics, with a branch explicitly called ‘experimental economics’, receives special attention.

After the recognition of the transition from observation to experiment, the notion of ‘experiment’ used in the social sciences is discussed. Thereafter, the focus is on the development of experiments in the social sciences, taking in particular Reinhard Selten’s contribution into account. His version of experimental economics touches important philosophical issues, both epistemological and methodological. Among them, prediction is a key notion asking for a suitable explication, especially in relation to the topics of accuracy and precision.

*Chapter 6: Ontological, epistemological, and methodological positions
(James Ladyman)*

Ontological issues in the philosophy of science may be specific to a particular special science, such as questions about the ontological status of biological species. They may also be more general, such as whether or not there are objective natural kinds and laws of nature, which requires in the first place explication of such notions. In the history of science ontological issues have often been of supreme importance; for example, whether or not atoms exist was a question that occupied many scientists in the nineteenth century. The particular epistemological problems raised by science mostly concern inductive inference, since it is widely accepted that substantive knowledge of the world cannot be obtained by deduction alone. The most fundamental of such problems is to explicate the relationship between theory and evidence, leading, for example, to the question whether the notion of ‘inference to the best explanation’ can be made sense of. Finally, methodology here means the theory of the scientific method. Is there a single such method for all the sciences, and if so what is it? How much should we expect the theory of the scientific method to help with the progress of science? Is the scientific method fixed, or does it change over time?

Chapter 7: Reduction, integration, and the unity of science: natural, behavioral, and social sciences and the humanities (William Bechtel and Andrew Hamilton)

Beginning with a brief review of historical notions of the unity of science, the chapter offers a ‘field guide’ to modern concepts of unity and reduction that starts with the theory reduction model of the logical positivists and then considers alternatives. In particular, the chapter discusses revisionist accounts of theory reduction, the best systems approach to unity, due to Phillip Kitcher, and the reasons why several thinkers — Suppes, Dupré, and Cartwright — find attempts at unity and reduction wrongheaded. With these notions and counter-notions in place, the chapter reviews arguments for integration instead of unification and for reduction in terms of mechanisms. Mechanisms involve lower-level parts and operations organized to yield higher-level effects and accounts of mechanistic explanation provide a novel perspective permitting integration without locating all causation at the lowest levels. Finally, case studies are offered of putative reduction and integration in thermal physics, molecular and developmental biology, archaeology, and linguistics.

Chapter 8: Logical, historical and computational approaches (Atocha Aliseda and Donald Gillies)

The chapter begins with the logical approach introduced by the Vienna Circle in the philosophy of science. The members of this circle confined the subject of

philosophy of science to the question of the justification of scientific theories. The dominance of the logical approach was challenged in the 1960s by the emergence (or perhaps better re-emergence) of the historical approach. This approach allowed the question of scientific discovery to enter philosophy of science. From the mid-1970s, the development of computers began to influence philosophy of science. Investigations in artificial intelligence led to the development of new logics, such as non-monotonic logics, which had been unknown to the Vienna Circle, and which enabled the logical approach to philosophy of science to be developed further. The study of machine learning shed new light on processes such as induction and abduction, enabling renewed explications, and it allowed the question of scientific discovery to be raised in a more formal manner.

Chapter 9: Demarcating science from nonscience (Martin Mahner)

The explication of the distinction between science and nonscience has been one of the main tasks in general philosophy of science. Most contemporary philosophers of science contend that there is no set of both necessary and sufficient criteria to demarcate the two. This chapter deals with the questions of (a) why we should nonetheless distinguish science from nonscience, and in particular from pseudoscience; (b) how we actually can distinguish science and nonscience even if there is no set of necessary and sufficient criteria; and (c) what the appropriate units of such demarcation are (e.g., epistemic fields, theories, methods). To this end, a comprehensive list of descriptive and normative science indicators, characterizing scientific epistemic fields, is proposed. This list allows for an analysis of various fields, such as mathematics, technology and humanities, as well as suspected pseudosciences, as to their actual status as a science (or at least a protoscience) or else as a nonscience, in particular a pseudoscience.

Chapter 10: History of philosophy of science (Friedrich Stadler)

(Post-)modern philosophy of science has been strongly influenced by the direct and indirect contributions of Logical Empiricism, that is, the Vienna Circle around Moritz Schlick and the Berlin Group around Hans Reichenbach, including its critics, Ludwig Wittgenstein and Karl Popper. Since the beginning of the 20th century we can reconstruct a long-term transfer, transformation and interaction of Central European philosophy of science to the Anglo-Saxon world: from *Wissenschaftslogik* (Carnap) to philosophy of science, and back to the (analytic) *Wissenschaftstheorie*. This significant development, brought on by the forced emigration of logical empiricists in the Nazi era, manifests the destruction of a creative network of philosophy of science as well as the intense interaction of scientific philosophy and philosophy of science in Central Europe, including the forgotten ‘French connection’ that existed with Pierre Duhem and Henri Poincaré, with the scientific community in Great Britain and North America from the 1930s to the 1960s. The latter were represented by neo-pragmatism and operationalism, which centered

around Percy W. Bridgman, Willard Van Orman Quine and Charles Morris as well as by linguistic and scientific philosophy of, notably, Bertrand Russell, Susan Stebbing, Frank P. Ramsey and Max Black.

A critical reconstruction of today's history of philosophy of science on the basis of exile studies and history of science, highlights this transatlantic movement and theory dynamics culminating in the long neglected re-transfer of analytic philosophy (of science) back to its roots of the 'third Vienna Circle', around Viktor Kraft, with Arthur Pap, Paul Feyerabend, and Wolfgang Stegmüller. Following the 'linguistic turn', the pragmatic and historical turns in recent philosophy of science (Quine and Kuhn) constitute an essential part of these developments in the period from Hot to Cold War.

4 ACKNOWLEDGEMENTS

On several occasions I had the opportunity to discuss the 'meta-explication of 'explication', for example, in Lisbon (2005), Opole (2006), and a couple of times in Groningen. I like to thank all discussants for their critical and constructive remarks. Moreover, I like to thank David Atkinson and Jeanne Peijnenburg for their detailed remarks on the final manuscript.

BIBLIOGRAPHY

- [Carnap, 1950] R. Carnap. On Explication, in: R. Carnap, *Logical foundations of probability*, University of Chicago Press, 1950. Second edition, 1963, pp. 2–8.
- [Carnap, 1966] R. Carnap. Three kinds of concepts in science, in: R. Carnap, *An introduction to the philosophy of science*, pp. 51–61. Basic Books, New York, 1966.
- [Hempel, 1945/1965] C. G. Hempel. Studies in the logic of confirmation, in: C. Hempel, *Aspects of scientific explanation*, pp. 3–46. The Free Press, London, 1965. First published in *Mind*, 1945.
- [Hempel, 1950/1966] C. G. Hempel. The empiricist criterion of meaning, in: *Logical Positivism*, edited by A.J. Ayer, pp. 108–126. New York: The Free Press. (Originally published in 1950.)
- [Hempel, 1952] C. G. Hempel. *Fundamentals of concept formation*, University of Chicago Press, Chicago, 1952.
- [Joyce, 2003] J. Joyce. Bayes' Theorem, *The Stanford Encyclopedia of Philosophy (Winter 2003 Edition)*, Edward N. Zalta (ed.), 2003. <http://plato.stanford.edu/archives/win2003/entries/bayes-theorem/>
- [Kememy and Oppenheim, 1952] J. G. Kemeny and P. Oppenheim. Degrees of factual support, *Philosophy of Science*, Vol. 19, 305–330, 1952.
- [Krajewski, 1977] W. Krajewski. *Correspondence principle and growth of science*, Reidel, Dordrecht, 1977.
- [Kuhn, 1962] T. Kuhn. *The structure of scientific revolutions*, University of Chicago Press, Chicago, 1962.
- [Kuipers, 1983] T. Kuipers. Non-inductive explication of two inductive intuitions, *The British Journal for the Philosophy of Science*, **34.3**, 209–23, 1963.
- [Kuipers, 2000] T. Kuipers. *From instrumentalism to constructive realism*, Kluwer, Dordrecht, 2000.
- [Kuipers, 2001] T. Kuipers. *Structures in science*, Kluwer, Dordrecht, 2001.
- [Kuipers, 2002/to appear] T. Kuipers. O dwóch rodzajach idealizacji i konkretyzacji. Przypadek aproksymacji prawdy, in: J. Brzezinski *et al.* (eds), *Odwaga Filozofowania. Leszkowi*

- Nowakowi w darze*. Wydawnictwo Fundacji Humaniora, Poznan, pp. 117-139, 2002. English version, Two types of idealization and concretization, to appear.
- [Laudan, 1977] L. Laudan. *Progress and its problems*, University of California Press, Berkeley, 1977.
- [Niiniluoto, 1987] I. Niiniluoto. *Truthlikeness*, Reidel, Dordrecht, 1987.
- [Nowak, 1980] L. Nowak. *The structure of idealization*, Reidel, Dordrecht, 1980.
- [Popper, 1963] K. R. Popper. *Conjectures and refutations*, Routledge and Kegan Paul, London, 1963.
- [Sober, 1998] E. Sober. Simplicity (in scientific theories), *Routledge Encyclopedia of Philosophy*, 1998.
- [Thagard, 1988] P. Thagard. *Computational philosophy of science*, MIT-press, Cambridge, 1988.